

Choreographic Buttons: Promoting Social Interaction through Human Movement and Clear Affordances

Andrew Webb, Andruid Kerne, Eunyee Koh, Pranesh Joshi, YoungJoo Park, Ross Graeber

Interface Ecology Lab
Department of Computer Science
Texas A&M University
College Station, TX 77843, USA

{awebb, andruid, eunyee}@cs.tamu.edu, {pranesh, rgraeber}@tamu.edu

ABSTRACT

We used human movement as the basis for designing a collaborative aesthetic design environment. Our intention was to promote social interaction and creative expression. We employed off-the-shelf computer vision technology. Movement became the basis for the choreography of gestures, the development of gesture recognition, and the development of imagery and visualization. We discovered that the design of clear affordances is no less important in movement-based than in mouse-based systems. Through an integrated and iterative design process, we developed a new type of affordance, the *choreographic button*, which integrates choreography, gesture recognition, and visual feedback. Jumping, a quick movement, and crouching, a sustained gesture, were choreographed to form a vocabulary that is personally expressive, and which also facilitates automatic recognition.

How can we evaluate socially motivated interactive systems? To create a context for evaluation, we held an integrated exhibition, party, and user study event. This mixing of events produced an engaging environment in which participants could choose to interact with each other, as well as with the design environment. We prepared a mouse-based version of the design environment, and compared how people experienced it with the movement-based system. Our study demonstrates that movement-based affordances promote social interaction.

Categories and Subject Descriptors

H.5.2 [User Interfaces]: User-centered design, Evaluation/methodology; J.5 [Arts and Humanities]: Performing arts.

General Terms

Design, Experimentation, Human Factors

Keywords

choreographic buttons, computer-supported cooperative play, gesture, movement, Laban notation, iterative design, mappings

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MM'06, October 23–27, 2006, Santa Barbara, California, USA.

Copyright 2006 ACM 1-59593-447-2/06/0010...\$5.00.

1. CONCEPT

Research on human computer interaction involving computer vision-based motion tracking needs to start with the body, and then consider technology. The body is the site of experience. We sense the environment, form understandings, and actuate responses. Movement is an actuated essence of people's everyday experiences. Movements articulate locomotion, and also convey information regarding emotions and intentions. A gesture is a linguistic form of movement that communicates without the use of vocal articulation or written language. Sign language is an example of a gestural system. Even outside of formal linguistic structure, gesture is a significant mechanism of human to human communication. Thus, gesture is social. Gesture can be utilized to develop expressive vocabularies that form a basis for the interaction between humans and machines. Choreography is the creation or specification of movement in order to affect a specific intention and convey meaning. In this research, human movement serves as the basis for human computer interaction, social interaction, and visual imagery. Expression through movement forms the basis of an interactive environment in which participants collaborate to create design. The design of clear interaction through affordances and feedback plays a key role.

2. INTRODUCTION

We develop a collaborative aesthetic design environment with a movement-based interface. This paper begins by describing the design process, and its results. The movement-based design environment consists of five structural components: (1) superimposed choreographic button duos; (2) a human Choreography Grid and a visual Imagery Grid; (3) a Movement Imagery Collection; (4) temporal visual structures; and (5) interactive affordances. We explicate about these components in a non-linear fashion corresponding to the non-linear structural relationships that emerged in the design process through which we developed their semantics, forms, and functions.

After describing the design process, we will develop an evaluation method that integrates exhibition, party, and user study events to create a context for investigating the social effects of the movement-based interface. To do this, we developed a mouse-based version of the design environment, and compared user experiences in the two scenarios. We combine quantitative and qualitative data to tell the story of the participants' experiences. The subsequent section reviews relevant prior work.

We close with discussion about the role of the body and movement in human computer interaction, the application of principles of interaction design in movement-based interfaces, and

methods for the evaluation of interactive systems designed to promote social interaction. As the traditional laboratory experiment is one context created by researchers, so it makes sense to develop new approaches to such context definition and development, to support the goals of new forms of research.

3. CHOREOGRAPHY AND RECOGNITION

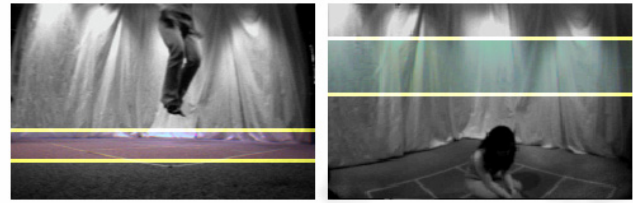
Our design process was focused on exploring possibilities for human expression, and the social function and utility of video motion tracking technologies [14], rather than developing new algorithms. What can be done with off-the-shelf video tracking? We utilized the Max signal processing and integration environment, along with the Jitter suite of video processing objects, and the Cyclops video analysis plug-in [2]. A program in this environment is known as a Max/Jitter “patch.” We developed gesture-recognition algorithms as a Max/Jitter patch.

We sought to choreograph gestures that were on the one hand expressive, and also not difficult to develop reliable recognition algorithms for. As we developed prototypes iteratively, we discovered that it is important to graphically convey the state of the gesture recognition process, in the course of its progress. Otherwise, the participant cannot clearly understand what state the environment is in, and what interactive movements make sense. Further, the physical space in which the participant can effect the environment also needs to be clearly defined, in order to make the possibilities for interaction clear. Norman associates with the term *affordance*, “the perceived and actual properties of a thing ... that determine how it can be used” [15]. Thus, we developed a tight binding between the demarcation of physical space, gesture and its recognition, and graphical representations of state. We call this affordance the *choreographic button*. Designing a choreographic button involves three interwoven stages: (1) defining (choreographing) and recognizing the movement; (2) designing graphical indicators that represent the state of the movement; and (3) mapping changes in the state of the movement to changes in the graphical affordances and the system.

3.1 Laban’s Schematic Movement Structures

Laban provides a structural approach to understanding movement and choreography [10], [11]. A person’s “movement arises from an inner volition that results in transference of the body or one of its limbs from one spatial position to another.” One aspect of movement corresponds to the shape of the body, while another addresses changes of the position of the body across space. We want to follow and understand the continuous creation of spatial impressions through the experience of movement. The unity of movement and space can be demonstrated by comparing single snapshots of the mind with each other, and showing that the natural order of their sequences and our natural orientation in space are based on similar principles.

Laban developed six elementary schematic structures of effort in correlation with the six fundamental directions in space: up and down, left and right, backward and forward [11]. We utilized the first of these duos. A tendency to go upward is correlated with lightness. An effort level of lightness, of losing strength, corresponds with the reaching upward to the point where the arm or the body prepares to relax and to fall back towards the ground. A tendency to go down is correlated with strength. A strong, firm movement always has at its source a vital connection with the



(a) floor camera view,
recognizes jumping

(b) upper camera view,
recognizes crouching

Figure 1: Profile view camera examples

stance. Laban said it is easy to feel that every strong movement is correlated to a foothold downwards. Moving left, right, backward, and forward are correlated to flexibility and a sense of spatial freedom. Quick movements correlate with sudden tensions and contractions, and slow movements with sustain and release.

3.2 Choreographing Recognizable Gestures

Based on a pair of Laban’s schematics of effort, we define a *choreographic button duo*: one button uses a quick jumping movement, while the other uses sustained crouching. In addition to the expressive nature of these movements, each was chosen because of the simplicity provided in computationally monitoring single-axis unidirectional movement (up and down). We recognize *jumping* movements with a single wide-angle video camera, and *crouching* movements with another such camera. Modern computational resources enabled our Max/Jitter patch to apply recognition algorithms to both video streams concurrently in real time using a single computer.

These two cameras are located in the corner to the northeast of the active choreographic area of physical space, providing a wide-angle profile view (Figure 2, right). Camera A is used to recognize jumping is located at floor level and provides a view of the feet (Figure 1a). By placing the camera parallel with the plane of the floor, only one algorithm is necessary to recognize jumping, because only a single rectangular area (junction between the floor and wall) needs to be analyzed regardless of the position of the feet along the floor. Camera B, which recognizes crouching, is located directly above and functions similarly to the floor camera, but with a view of the upper torso (Figure 1b). A curtain running just outside the south and west edges of the active choreographic area provides a static backdrop to reduce noise for motion tracking.

This recognition mechanism is sufficient to support a single choreographic button duo corresponding to a horizontal area that spans the range of the cameras. Lateral spatial resolution requires employing a third camera to track the position of participants. The camera (Camera C) that differentiates lateral positions is placed overhead, facing the floor, orthogonal to the other cameras. Lateral space can be discretized. Upon subdividing the feed from the overhead camera into discrete regions (cells), a choreographic button duo is assigned to each. By putting together the profile and overhead views, we are able to recognize distinct movement forms, in association with specific lateral positions.

4. A TALE OF 2 GRIDS

We divided a square grid, six feet on a side, into cells corresponding to a set of discrete lateral positions for human interaction. The size of each of cell, two feet by two feet, corresponds approximately to the area on the floor occupied by a

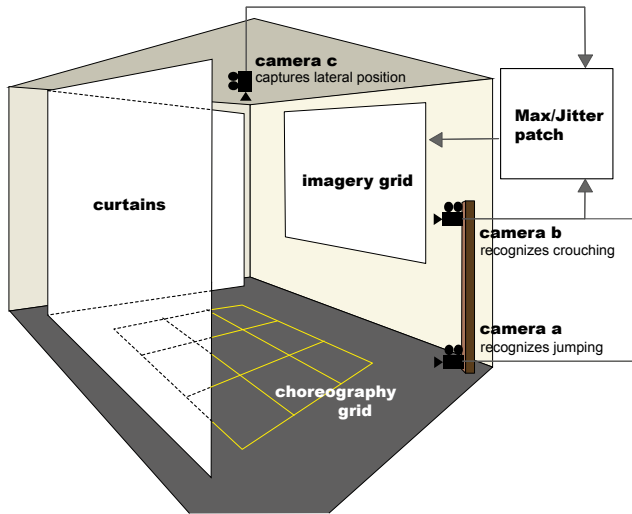


Figure 2: System Diagram: Choreography Grid, Imagery Grid, and their relationships to the space and technology.

quiescent human. The nine choreographic button duos of our movement-based design environment are manifested not just in physical space, but also in an associated visual space. The grid in the physical space, where gesture recognition for human computer interaction takes place, is called the Choreography Grid, while the projected space of visualization and feedback is the Imagery Grid (Figure 2). The two grids are characterized by a direct one-to-one positional mapping of cells. In other words, the top-left corner of the Choreography Grid maps to the top-left corner of the Imagery Grid. The consistent visual language and direct spatial mappings are easy to understand for participants

4.1 Choreography Grid

The boundaries for the cells of the Choreography Grid are marked out on the floor of the physical space with yellow lines. As a person walks from cell to cell in the Choreography Grid, selection of the corresponding cell in the Imagery Grid is indicated by a yellow box. To avoid confusion when a participant overlaps multiple cells, the Max/Jitter patch considers a participant to have moved to a new cell only when the majority of the participant's body is positioned there.

4.2 Imagery Grid

The Imagery Grid (Figure 3) projected on the forward facing wall is the arrangement of the imagery states for the choreographic button duos. The 3x3 grid structure provides a collage creation structure with degrees of freedom designed to support an experience that is aesthetically evocative and creatively open while avoiding clutter and expressive confusion. The discretized grid combines simplicity and complexity to promote design.

5. MOVEMENT IMAGERY COLLECTION

Visual representations rendered in each cell of the Imagery Grid are based on the Movement Imagery Collection, a set of 29 images that we chose in order to represent forms of movement in a physical or emotional sense, developing a movement vocabulary based on the gestures of the choreographic buttons: jumping and crouching. The subjects of these images are diverse; they include a child playfully jumping on a small merry-go-round, up and down views of staircases, and an emotionally distressed youth

crouched on the ground. We use this narrative imagery to amplify the movement mappings of our choreographic buttons. The images were selected with a curatorial sense to form a collection of building blocks, so that any juxtaposition will be visually coherent and semiotically evocative. By using images with these relationships, our intention is to help participants without previous knowledge of how the design environment works to derive the appropriate movements. These images also provide an emotional narrative for the participant to explore. Many of the images convey more than just the movement of jumping or crouching, but also an emotional experience that is either introverted or extroverted. They represent heightened states that reflect personal processes of transformation. The intention in providing emotional narratives that reflect the affordances is to immerse the participant in the experience of the design environment. As well, jumping and crouching repeatedly is physically demanding. The effort creates a somatic form of immersion. We intend to create a reflective, engaging experience in which movement transforms imagery and the experience of the self. Through physical and emotional engagement, we seek to motivate the experience of the design environment, itself.

6. TEMPORAL VISUAL STRUCTURES

We designed three temporal visual structures for selecting elements from the Movement Imagery Collection and rendering them: cross-fading, fast-forwarding, and still. Each of these structures consists of rules for selecting an element, an option for compositing with the previous element, and an option for sequencing over time.

The first and default structure is *cross-fading*. When a grid cell is cross-fading, the visualization continuously transitions between pairs of images from the Movement Imagery Collection. Each pair consists of a source image and a destination image. Each successive destination image is selected randomly. Cross-fading progressively renders source and destination images with inverse levels of translucence (Figure 4), as per these equations:

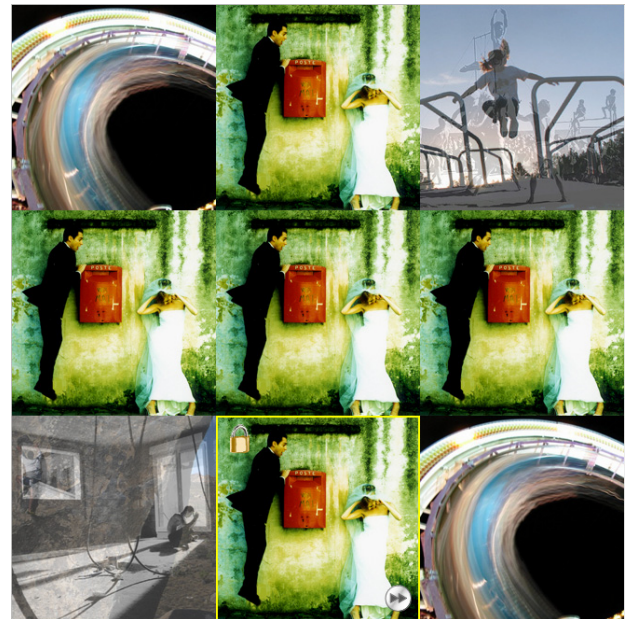


Figure 3: Imagery Grid Example. The bottom center square is selected and locked.



Figure 4: Temporal Cross-fading Example: Excerpt from the sequence of frames that would appear successively in a single grid cell; the transition from a source to a destination image.

$$\alpha_{source} : [0,1]$$

$$\alpha_{dest} = 1 - \alpha_{source}$$

Following each cross-fade, there is a five-second pause before the next cross-fade occurs with a new destination image. The pause provides a participant with time to discern individual images in the collection, as well as their blended states.

At initialization, every cell in the Imagery Grid is invoking the cross-fading temporal structure. The initialization times of the cross-fading structures in different cells are offset, in order to desynchronize the fades. This conveys a sense of mutual independence, rather than synchrony.

The second temporal visual structure is *fast-forwarding*. When this structure is invoked, the imagery transitions between discrete images from the Movement Imagery Collection in a fixed sequence. There is no compositing. A pause between transitions allows the participant time to observe each successive image, and make decisions about whether to select it with a movement.

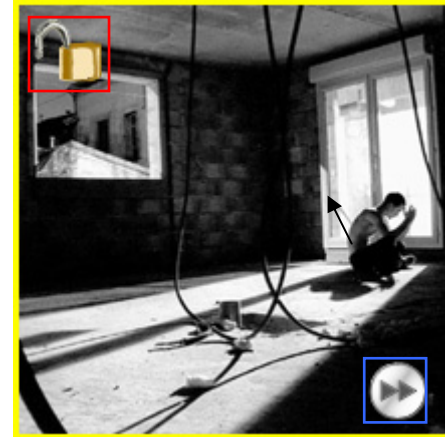
The simplest structure is *still*. This is the identity element. In the still structure, the imagery visualization is fixed to a single image from the Movement Imagery Collection. The still temporal visual structure does not transition between images.

7. AFFORDANCES AND MAPPINGS

The participant moves across the space of the Choreography Grid in order to begin interacting with the environment. How does s/he know what interactions are possible? Creating clear affordances is essential to the design process for any interactive interface [15]. Designers must ensure that the participant/user knows the current state of the interaction, to avoid causing confusion and frustration. The same is true about informing the participant/user about where to “press.” Norman explains the manner in which humans interact with systems through an Execution-Evaluation framework [15]. An interaction can be divided into two phases – the Execution phase in which a participant performs actions on the world (via components of a system) and the Evaluation phase in which the participant senses the state of the world, to evaluate the results of his actions. In order to minimize the gulfs of execution and evaluation, the system design should help the participant build a correct conceptual model of the system, make essential parts visible, provide memory aids to the participant, provide clear feedback, and accommodate errors [15]. These principles apply to movement-based, as well as mouse-based interfaces.

We choreographed the gestures of quick jumping and sustained crouching as the basis for each button duo. Each duo is associated with a square cell of physical space in the Choreography Grid, and a corresponding cell in the projected visual space of the Imagery Grid. In the physical space of the Choreography Grid, while the choreographic buttons of a duo are superimposed laterally, their positions are differentiated along the vertical axis.

Icon for locked or unlocked



Icon for fast-forwarding

Figure 5: Affordances Example. The button duo is selected and locked. Fast-forward is not activated.





























Table 1 defines the mappings by specifying the state machine of associated movements, affordances, temporal visual structures, and transition logics. The remainder of this section describes the rules through which these mappings are manifested.

Selection of a choreographic button duo occurs when a participant moves laterally into a particular Choreography Grid cell. This is indicated in the visual space of the Imagery Grid by a yellow selection border that frames the corresponding cell. Selection also results in the presentation of indicators reflecting the state of each button in the duo. The button state icons are positioned vertically, corresponding to the movements required to trigger each of them. The jump button state icon is located in a top corner since upward movement triggers it. The crouch button state icon is located in a bottom corner because downward movement triggers it. Thus, the crouching button affordance is activated by moving down to remove oneself from the upper area; while jumping up out of the lower area activates the other button.

In developing the design for iconic visual feedback and mappings to temporal visual structures (system states) once a choreographic button duo is selected, we were guided by physical constraints and the characteristics of movement. Jumping is quick and sudden, and thus well suited for a momentary state-changing operation. Locking is a toggle operation that transpires instantaneously. Thus, jumping is mapped to the toggle of a lock state. Locking is indicated by a padlock that is open when not activated or closed when activated. Once it has been locked, a choreographic button duo invokes the still temporal structure, whether it is actively selected by the participant, or not. Locking also inhibits the affects of crouching. Crouching is a sustained gesture that can last for an indefinite amount of time. When a button duo is unlocked, crouching is mapped to the sustained selection of the fast-forwarding temporal visual structure. The fast-forwarding icon is a circular button with double forward arrows that is silver when not activated and blue when activated (Figure 5).

After spending time interacting with the button duo within a Choreography Grid cell, the participant will eventually move laterally. The cell becomes unselected, and the yellow selection border is removed in the Imagery Grid. When a grid cell enters

Table 1: Temporal Structure and Affordance Mappings Based on Movements

	Before Movement		Movement	After Movement	
	Temporal Structure	Affordances		Temporal Structure	Change in Affordances
unselected	cross-fading		walk into	still	yellow selection box open lock    non-activated fast-forward
	still (locked)		walk into	still	yellow selection box closed lock  
selected and unlocked	still	yellow selection box open lock    non-activated fast-forward	walk out	cross-fading	all removed
	still		jump	still	closed lock icon  
	still		crouch	fast-forward	activated fast-forward icon   
selected and locked	still	yellow selection box closed lock  	walk out	still	all removed
	still		jump	still	open lock icon   
	still		crouch	still	no change  
	still		stand up	still	no change  
	fast-forward	activated fast-forward   	stand up	still	non-activated fast-forward   

the unselected state, it either invokes the cross-fading temporal visual structure if it is unlocked or the still structure if it is locked.

The interactive semantics of locking allows a person to create a collage that is entirely static (all cells locked) or a heterogeneous dynamic composition in which some cells sustain cross-fading, while others are locked to still. This enables the participant/user to create patterns (see example, Figure 3). Changes made in the grids persist even after a person leaves the design environment space. A subsequent participant can enter the space and make changes to the previous creation, consequently making the new collage a collaborative work between the current participant and prior participants. Figure 6 shows an example of the Imagery Grid being used by a participant.

8. SCENARIO

Take for example this scenario of a participant, Sara, using the design environment. Initially, all choreographic button duos are unlocked, unselected, and invoking the cross-fading temporal visual structure. Sara walks into the Choreography Grid at the bottom right cell (SE corner). Immediately, the bottom right cell of the Imagery Grid switches to the still visual temporal structure. The yellow selection border and unlocked and non-activated fast-forward icons appear in the Imagery Grid cell.

Sara notices an interesting image in the Imagery Grid cell directly to her left. She sidesteps left into the bottom middle cell in the

Choreography Grid. The state indicators in the bottom right cell are removed, and the cell switches to the cross-fading temporal structure. Simultaneously, the bottom middle cell of the Imagery Grid switches to the still temporal structure and the selection border and two state indicator icons are visualized in the cell. Sara decides to keep the image, so she jumps directly up. The unlock icon switches to the lock icon, indicating that the choreographic button duo for the bottom middle cell is now locked. The fast-forward icon is removed since the fast-forward temporal structure is unreachable from the locked state.

Sara notices an interesting image that is fading out in the center cell. She steps forward into the center cell. Since the bottom middle cell is locked, it remains in the still temporal structure, but the state indicators are removed as the cell is no longer selected. The center cell is selected and switches to the still temporal structure with state icons appearing. Sara just missed the image she saw. The image shown in the center Imagery Grid cell is the one that was being cross-faded into. In order to obtain the image Sara originally desired, she must fast-forward through the images. Sara crouches causing the temporal structure to change to fast-forwarding. The fast-forwarding icon switches to the activated yellow version. Sara remains crouched until the image she wanted appears. She stands up. The temporal structure switches to still and the fast-forward icon switches back to the non-activated icon. She jumps to lock the cell.

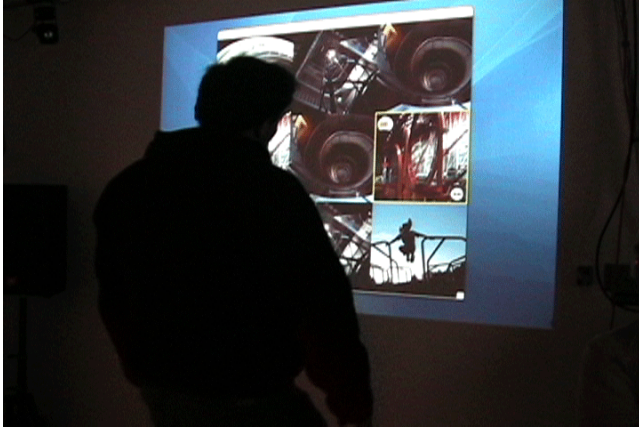


Figure 6: The Imagery Grid in use by an exhibition / party / user study participant.

Sara continues interacting with the design environment. Another participant, Tom, walks near the Choreography Grid and observes Sara's interactions. He comments to her about how he enjoys her current Imagery Grid composition. He suggests changing the top left cell to a different image. She tries his suggestion by crouching in the top left cell of the Choreography Grid to fast-forward to his suggested image. While the fast-forwarding is occurring, a conversation arises between the two about the physical interactions with the design environment. His suggested image appears, but Sara is preoccupied with the conversation and misses it. She quickly stands up just as the image changes. The two have a laugh about the experience. Sara eventually finishes interacting with the design environment. She exits the Choreography Grid.

9. EVALUATION

9.1 Study Design

We hypothesized that our design environment with choreographic buttons would afford and facilitate social interaction. To investigate this hypothesis, we held an exhibition / party / user study. Amalgamating these three different types of events allowed us to use HCI evaluation techniques in a realistic (semi-controlled) environment while giving the participant the opportunity for an experience less like that of a typical experimental subject, and more like a participant at a social art event. Signing a study consent form acted as a ticket to the event. Food was served. Popular music was played, sometimes at high volumes.

We need to discover evaluation methods that match activity-oriented intentions for interactive system design, such as to promote movement, play, and social interaction. Crabtree considers a breaching experiment to be a design deployment scenario in which participants are provoked with unexpected new technological systems [1]. There is no prior context of use that is being supported. Instead, novel systems are interjected into new contexts with participants who are not anticipating them, resulting in new practices. The present research deploys the choreographic buttons grid system design environment in the context of the exhibition / party / user study event with the intention of provoking physical activity and social interaction, and developing understanding of emerging practices of use.

In order to structure the evaluation process, we wanted to be able to compare two "usage" scenarios. Thus, we constructed a mouse-

based version of the aesthetic design environment. It is important to understand that this design was a retrofit of the environment created through the design process described above, in which movement was integral. It uses the same Movement Imagery Collection, Imagery Grid, temporal visual structures, and iconography as the movement-based design environment, but the interface is triggered with a mouse instead of with movements of the body. Thus, there is no Choreography Grid in this version. In order to select a square, a participant places the mouse cursor over the desired square. Locking/unlocking a square entails clicking the padlock icon for that square. Fast-forwarding a square is accomplished by holding the mouse button down over the fast-forward icon and releasing the mouse button when complete. Fast-forward was programmed to run more rapidly in the mouse-based version, since the action to stop fast-forward is a much quicker reflex – lifting a finger versus lifting the whole body.

The two systems were placed in the same room at opposite ends. The visualization for each system was projected at the same size onto a different wall. For the mouse-based version, a small school desk with a wireless mouse was placed several feet away from and facing the projection.

Participants were asked to spend time using each system. They were free to choose which system they used first. However, only one person could use each system at a time, which provided an incentive for using whichever one had the shortest wait.

Video recording served as a means for collecting qualitative data such as the social interactions between a participant using one of the systems and other participants. A tripod was placed near a central location in the room, allowing the cameraperson to smoothly switch between the two systems. The cameraperson took the camera off the tripod occasionally to acquire handheld shots of interactions that could not be seen from the tripod.

Each participant was asked to fill out a post-questionnaire. The post-questionnaire contained several structured questions asking for comparisons between the two systems, such as which is more engaging, more entertaining to watch others use, and more involving social interaction. Possible answers were either of the two systems or that both are equal. Following each comparison question was an open-ended question asking the participant to explain how and why s/he selected her answer.

The study team conducted a debriefing session following the end of the second session. This debriefing session was documented by videotape to help in later analysis. In the debriefing session, we discussed what we learned from the study. As well, we commented on which parts of the study went smoothly and which parts did not.

Participants were recruited by flyers, email, and word of mouth. Participants were invited from a range of academic areas (e.g. computer science, art, design). Twenty-two people responded. Nineteen post-questionnaires were filled out. The study was divided up into two sessions each lasting a little under two hours.

9.2 Results

We were unsure to what extent the mixing of the party and user study modes of social interaction would work. User studies are typically completely structured, where parties involve unstructured spontaneous social interaction. A fear was that the socializing would be stiff, due to the unexpected juxtaposition of processes. In fact, the participants seemed to be quite comfortable

with the mix. They voluntarily took responsibility for engaging in the structured activities of using the two systems, and answering the questionnaire. They also were happy to eat food, and engage in free form conversation. The integrated form of the event elicited discussion from participants about the interaction design issues of choreographic buttons and the Imagery Grid.

9.2.1 Quantitative Experiential Data

The comparison questions from the post-questionnaire provided with quantitative experiential data (Figure 7). Participants effectively distributed themselves as to which system they used first. The movement-based version had three more participants use it first than the mouse-based version. Nine participants found the movement-based version of the system to be more engaging, while eight participants found the mouse-based version more engaging. The remaining two participants found both systems equally engaging. There is no significant difference between the movement and mouse based versions in terms of engagement [$\chi^2(2) = 4.53, p = 0.104$]. The next question asked which version of interaction is more entertaining to watch others use? Nearly all the participants (18) found the movement-based version more entertaining to watch as other participants used it; the result is statistically significant [$\chi^2(1) = 15.21, p < 0.001$]. Participants also found a difference between social interactions that occurred while using the different systems [$\chi^2(1) = 14.22, p < 0.001$]. Most (15) felt that the movement-based version involved more social interaction. However, a few participants (3) thought the mouse-based version involved more social interaction. This result shows that the participants were more socially involved and interacted more with each other using the movement-based version than the mouse-based version [$\chi^2(1) = 8.00, p < 0.01$]. Most participants (13) felt emotional connections with the images displayed, but some (5) did not. The result borders on statistical significance [$\chi^2(1) = 3.56, p = 0.059$].

9.2.2 Qualitative Data

The open-ended questions from the post-questionnaire and video taken during the study provided us with qualitative data. Participants who found the movement-based version more engaging said that it involved movement of their entire body, produced good exercise, or was fun to play with like a game.

"I felt a sense of connection to my movement-based composition (as I had literally invested sweat). It felt exciting to move my entire body and have a system respond."

The other half of the participants found the mouse-based version more engaging for its ease of use, interface familiarity, and limited requirements for physical engagement. The mouse-based version seemed to provide a comfortable interface for them. Sitting down at a desk helped these participants focus more on the images and composing.

"Holding the mouse with my hand gave me more 'control' over what I was doing. I was able to look at the image with more detail."

As was also demonstrated by the quantitative results, participants found the movement-based version socially engaging when they were in the role of observer, because of the performance created by another person using the system.

"Clicking is not enjoyable to observe, but moving seems non-deterministic in that an anomaly could occur that could influence

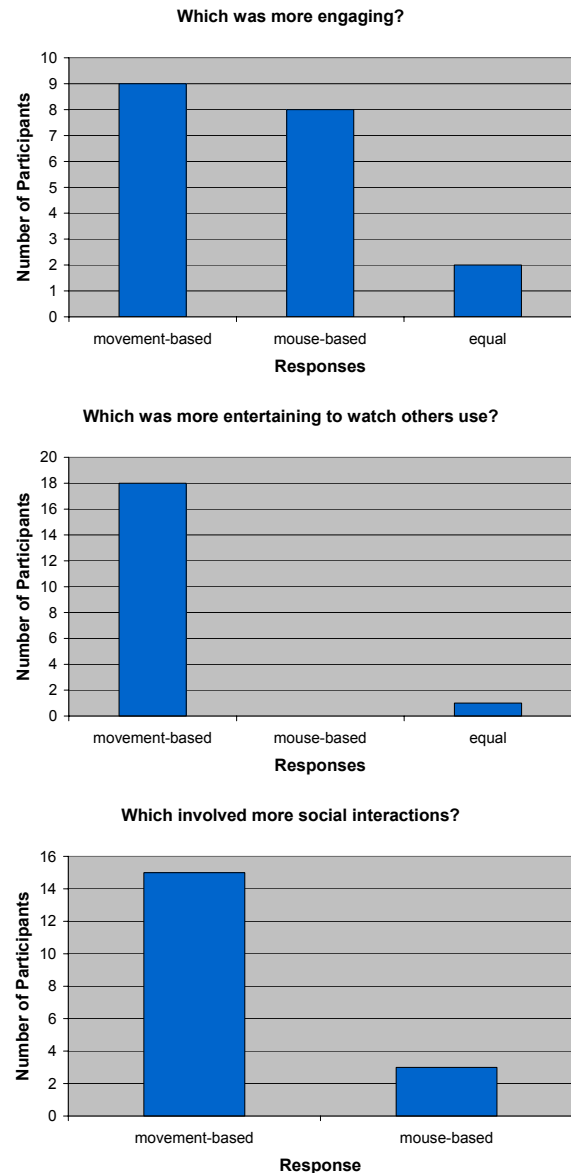


Figure 7: Quantitative experiential results.

the composition in interesting ways. People asked questions and added their opinion when I moved, but not when I clicked. The movement seems to (socially) break the ice. It is compelling, not ordinary (as the click is)."

Two of the participants noted entertainment value in the movement-based system and its environment for the study due to similarities seen with dancing games such as Dance Dance Revolution. Computer-supported cooperative play in the movement-based design environment was reinforced through the system's collaborative and ludic nature as expressed by these participants.

"It's funny to see people jumping and concentrated in looking at the screen. It seems more natural to talk and interrupt people in the movement-based than in the mouse-based."

Several participants (5) discussed between each other about the novelty of the movement-based system, which helped incur more

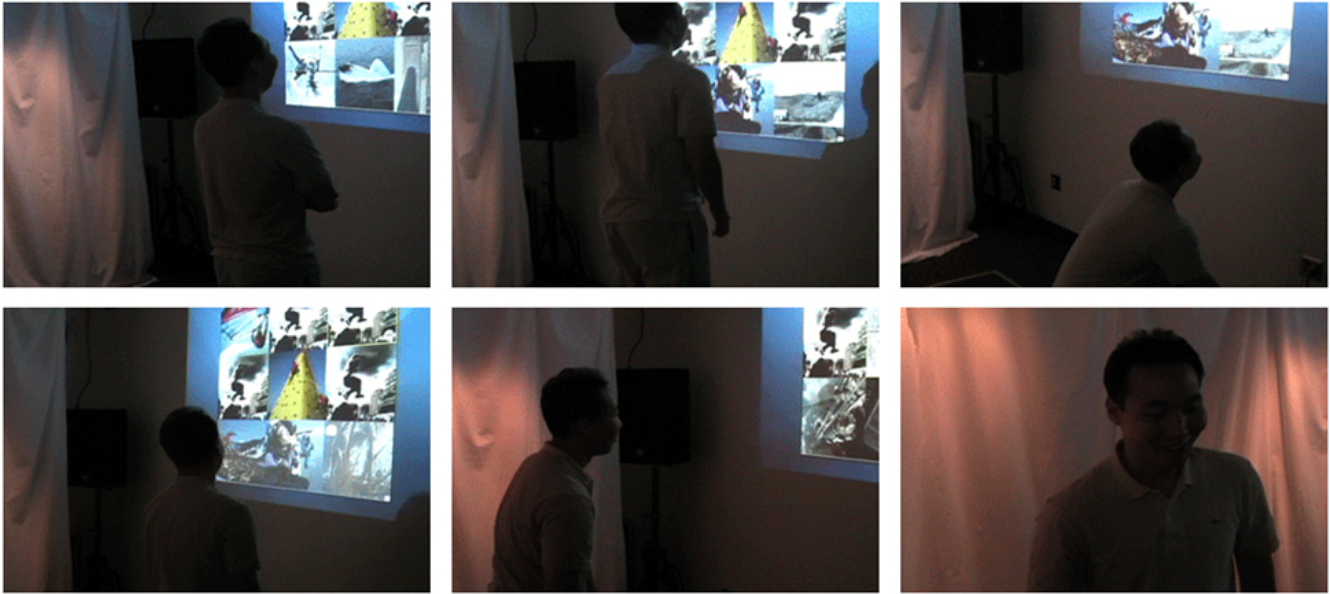


Figure 8: Participant from user study interacting with movement-based system

social interactions between participants. As well, the novelty of the system itself increased social interactions between participants.

“Moving is very different. This difference causes interesting questions to arise. It feels easier to question and comment about someone who is rigid and socially stiff. The vulnerability of the composer makes initiating conversation easier.”

One participant felt isolated when using the movement-based system.

“On the movement-based, you, the user, were isolated from the large group; whereas, with the mouse-based, others were around to give feedback about the collage.”

Two participants noted that the movement-based system lent itself more to a collaborative composition between the participant using the system and those watching; the mouse-based system lent itself to individual composition. Most of the participants found a difference between their interactions with the two systems. Three participants found the movement-based high energy, while the mouse-based was efficient for viewing images and making compositions. Five other participants found the movement-based captivating versus the mouse-based being ordinary. And still, five others found the movement-based more fun versus the mouse-based being more for composing.

“Night and day. Mouse-based seemed ordinary & therefore somewhat boring; the movement kept me on my toes (literally) through the entire process of authoring.”

Another collection of qualitative data came from the video captured during the two sessions. The video was filled with examples of social interactions between participants related to the use of either of the two systems. In several instances, a participant asked for compositional suggestions while using the movement-based system. As well, other participants observing the use of the movement-based system made comments about the composition between each other and the participant being observed. For example, one participant asked another, “Is everything locked on

it now?” The participant responded, “Yeah, I wanted that thing in the corner to be a circle too, but it moved on me before I could lock it.” In another example, one participant questioned another, “What are you doing?” The response was, “I don’t know,” followed by laughter between the two. The video contained instances from both the movement-based system and the mouse-based system where a participant was explaining his/her process in composing to others nearby.

“Right now I’m playing with color more than shape because I was playing with shape versus color earlier.”

Other conversations among participants involved discussing what purpose the movement-based system served, how the tracking works, and how novel the system was. For instance, one participant asked another, “What is the purpose of the system?” The other participant responded, “Artistic expression,” and another nearby participant chimed in with, “and potentially exercise.”

10. RELATED WORK

This work is related to prior work in a number of ways. Some work is relevant due to its use of body-based interfaces for aesthetic design responsive environments and other interactive art installations. Other work is similar in terms of its reflective design process, and evaluation methodologies that reach beyond usability and efficiency measures.

10.1 Body-Based Interactive Art Installations

Our movement-based design environment has attributes in common with prior interactive art installations, which constitute forms of physically *responsive environments*. Krueger introduced the concept of a responsive environment in which a computer perceives the actions of those who enter and responds intelligently through complex visual and auditory displays [8]. In Videoplace [9], Krueger designed a virtual environment that connects people in remote locations in a common visual experience, allowing them to interact through the video medium. Mandala extended Videoplace in that it represented participants with their video

images, while incorporating 2D objects and characters in the scene [19]. Interactive Video Environment is known as the first attempt to extend the virtual space, objects, and interactions to 3D [3][12]. Video images of the participants captured by multiple video cameras are composited to form an image of a 3D Virtual Environment inhabited by 3D objects and agents. The compositing system enables participants to interact within the Virtual Environment using full-body motions and gestures. The function of visual imagery in our design environment is similar to these works in that it functions to afford communication, art, and entertainment; however in our work, participants interact with visual images, affordances, and each other to create their own solo and collaborative interactive art forms. In addition, our work focuses on designing participatory choreographic affordances.

“Puccini Set Designer” is an installation utilizing a computer vision interface to enhance the viewer’s experience of the work of Puccini [18]. The purpose of this installation was to accentuate the viewer’s sense of participation in the process of experiencing art. Five “surrogate icons” painted on the floor in the physical space are termed the “Carpet of Light.” They represent five sets from Puccini’s opera. When a person steps on any one of the surrogates, the vertical screen in front of him displays a set of dynamic pictures. By pointing to different parts of this picture, a participant can trigger a particular subset in the Puccini’s opera. Both the Choreography Grid and the “Carpet of Light” serve similar roles (spatial arrangement of buttons) in their respective systems. The choreographic button and the painted surrogate icon also play similar roles while requiring different movements. We utilize whole body-based gesture rather than hand gesture and create a correspondence of spatial mappings between physical and graphical space. As well, in our responsive design environment the participant not only experiences art but also creates art.

As our design environment captures body-based movements, “living-room” considers the position of the participant and the orientation of the participant’s gaze to manipulate an augmented reality system [4]. “living-room” explores interactive, space-related aspects of augmented reality in art and design fields like architecture, interior design, and scenographic design. Changing contrast, color, light saturation and other parameters of the recorded image allow the participant to transform the look of the real space and make it more similar to the synthetic world.

Schiphorst applied Laban in the design of a touch input system utilizing garments made from directionally conductive fabric, instead of video tracking, to recognize gestures [17]. The wearer uses this fabric as an interface to select interaction modes that direct data between networked garments in a wearable art installation. Through gestures and the garment interface, the wearer can control, select, and explore the interpersonal exchange of physiological data. These connections form the basis of human to individual human and human to group social interaction. Our design environment also uses motion to generate experiential data and interactive mappings, and promotes social interaction.

Others have also created socially oriented aesthetic design environments that utilize projection and non-mouse based interfaces. Huang and Waldvogel describe the history of how the wall and its surface (wallpaper), in concert with technologies, have functioned as devices for social and spatial interaction [7]. Word Map Wallpaper uses microphones to capture pieces of conversations and projects these words on a nearby wall [7]. As a

person approaches the wall, the words begin to move toward her. Our design environment also uses the wall surface as a visualization medium and physical presence as an interface.

10.2 Design Goals and Evaluation Methodologies

We are building on a foundation of prior work that explores the use of HCI evaluations for helping artists to assess the efficacy of interactive installations. Crabtree’s breaching experiments, in which new contexts and practices of use are developed around new technologies, have been described above [1]. Höök *et al* raise important issues about how such evaluations must deal with factors beyond usability [6]. They offer appropriate situated methods for evaluating aspects of The Influencing Machine, according to the artists’ stated goals. Construction of the context of evaluation, the social process, and the data gathering mechanisms are all significant parts of such evaluation processes. We also constructed an evaluation process in terms of our goals as artists, as well as interaction designers by providing the socially interactive environment in a real world setting.

Gaver *et al.* present an interactive coffee table that uses the arrangement of weights on the surface to control the display of aerial photographs shown through a viewing hole in table’s surface [5]. The “Drift Table” was designed for ludic activities. By *ludic*, they mean, “... activities motivated by curiosity, exploration, and reflection, rather than externally defined tasks.” Our movement-based design environment was also designed for ludic engagement. We selected movements like jumping to encourage playful interaction. The system’s design possibilities are more oriented toward engaging in design processes, in concert with companions, than producing particular products. Similar to our evaluation methodology, [5] also considers social interactions with the Drift Table as an essential component for evaluation.

11. DISCUSSION

The experience of human computer interaction should be no less than other experiences in human life. As Merleau-Ponty develops, the sensation is the unit of experience, and the body is the site of sensation [13]. We humans are physical creatures. Thus, to bring human computer interaction onto a par with the rest of life, we need to discover new modalities that directly involve the body. Choreography is the field that has developed knowledge of expressive human movement. In framing this research, we set out to involve principles of choreography, of human movement, of effort and shape, in interface and experience design.

Computer-vision based motion-tracking technology served as the basis for recognizing human movement. Expressive movements and imagery were co-designed with gesture recognition algorithms. The role of technology was essential, yet not central in the interface design process. Further, while we used choreographic principles to design movement, we also needed interface design principles to develop a coherent experience for participant/users. The need for perceptible affordances, consistency, feedback, and clear mappings was if anything intensified, rather than obviated, by the involvement of computer vision. Choreographic buttons accomplish this with clear mappings from lateral physical space (the Choreography Grid) to visual space (the Imagery Grid), and likewise with clear mappings along the vertical axis, with jumping and the lock icon on top, and crouching and fast-forward iconography below. The consistency

of the graphical indicators, and the actuating motions resulted in a system design without usability issues. People learned immediately how to control the system. Choreographic buttons were found to produce a clear, intelligible user experience. The association of movement and visual state feedback was effective.

The visual and temporal design of the environment was developed in the grounded context of its choreography. Visual, interactive, and structural forms were produced about movement, in the context of movement. When the context of interaction was shifted from the movement-based to the mouse-based interface, these forms were found to still produce an effective experience.

One of our primary goals for the choreographic buttons system was to promote human to human social interaction. This manifests a value system in which developing human social relationships is considered worthwhile. As the theater director and performance studies scholar Richard Schechner said, "Process -- a term used often in environmental theater -- means 'getting there' rather than 'getting there,' emphasis on the doing, not the done... to be alive to the here and now, to express oneself here and now. What an immense risk that is! Those who love products value things and make things of all living beings. Those who love process value living and make living beings of all things" [16]. This research represents an activity-based rather than task-based, a process rather than product oriented goal structure.

One form of social interaction that we observed involved the use of the choreographic buttons grid system to create patterns. In the midst of each session of the exhibition / party / user study, one participant would create a clear pattern, using repetition and symmetry (see example, Figure 3). Other participants would comment on this. In some cases, subsequent participants would make new patterns that were somehow derivative of or related to the previous one. Thus, a chain of art works was created in relation to one another. In this way, social interaction developed through two stages, based on the structures of the collaborative aesthetic design environment. The grids themselves afforded the creation of patterns; the movements and the display afforded human to human social interaction, in which the grid-based patterns functioned as vocabulary. The result was the emergence of design style. Future research can further develop this paradigm.

How to evaluate an activity-oriented interactive system designed to promote social interaction and ludic engagement, rather than for the performance of specific tasks, is not straightforward. The feedback and observations garnered through the present study demonstrate a process through which meaningful and valuable data about such a system can be collected through the intentional mixing of socially engaging events and HCI evaluation practices. Specifically, an effective integration of exhibition, party, and user study events serves as a context for the evaluation of how a design environment promotes social interaction. The unstructured nature of the social events did not interfere with the procedures of our user study. The switching of context from "Hey, how's it going?" to "Please feel out this questionnaire," was accomplished seamlessly. Thus, we have developed a new methodology for the evaluation of systems designed to promote social engagement. We are creating a context, rather than depending on contexts that already exist. In the resulting breaching experiment context, participants are provoked with novel systems. More typical laboratory experiments take other approaches to manufacturing context.

Quantitative and qualitative data paint a consistent portrait: compared to a mouse-based interface, choreographic buttons promote social engagement. They got people who are not artists involved in making art together. They got people who are not dancers involved in moving in public. The reciprocity of turn taking, of the exchange of roles, seems to be a factor in this, along with the pure physicality of the interaction. Clear affordance design contributed to the efficacy of the movement-based interface.

12. REFERENCES

- [1] Crabtree, A. Designing in the Absence of Practice: Breaching Experiments. *Proc DIS 2004*, 59-68.
- [2] Cycling '74. Cycling '74: Max/MSP for Mac and Windows. <http://www.cycling74.com/products/maxmsp.html>. 2004.
- [3] Darrell, T., Maes, P., Blumberg, B., Pentland, A. P., A Novel Environment for Situated Vision and Behavior, Springer Series In Perception Engineering, 1995, 319-331.
- [4] Galantay, R., Torpus, J., Engeli, M., "living-room" Interactive, Space-Oriented Augmented Reality, *Proc MM 2004*, 64-71.
- [5] Gaver, W., W., Bowers, J., Broucher, A., Gellerson, H., Pennington, S., Schmidt, A., et al, The Drift Table: Designing for Ludic Engagement, *Proc CHI 2004*, 885-900.
- [6] Höök, K., Sengers, P., Andersson, G., Sense and Sensibility: Evaluation and Interactive Art, *Proc CHI 2003*, 241-248.
- [7] Huang, J., Waldvogel, M., Interactive Wallpaper, *Proc SIGGRAPH 2005 Art and Animation*, 172-176.
- [8] Krueger, M., Responsive Environments, From *AFIPS 46 National Computer Conf Proc*, 423-33, 1977
- [9] Krueger, M. *Artificial Reality 2*, Addison Wesley, 1990.
- [10] Laban, R., Lawrence, F. C., *Effort: Economy of Human Movement*, 2nd edition, MacDonald and Evans, 1973.
- [11] Laban, R., *The Language of Movement: A Guidebook to Choreutics*, Plays Inc., First American edition, 1974.
- [12] Maes, P., Alive: An Artificial Life Interactive Video Environment, *Proc. SIGGRAPH 1993 (Visual)*, 189-190.
- [13] Merleau-Ponty, M., *The Phenomenology of Perception*, New York: Routledge, 1995.
- [14] Moeslund, T.B., Granum, E., A Survey of Computer Vision-Based Human Motion Capture, *Computer Vision and Image Understanding* 81, 231-268 (2001).
- [15] Norman, D., *Design of Everyday Things*, Basic Books, 2002.
- [16] Schechner, R., *Environmental Theater*, NY: Applause, 1994.
- [17] Schiphorst, T., Jaffe, N., Lovell, R., Threads of Recognition: Using Touch as Input with Directionally Conductive Fabric, *Proc CHI 2005*, 2234-2238.
- [18] Sparacino, F., "Scenographies of the Past and Museums of the Future: From the Wunderkammer to Body-Driven Interactive Narrative Spaces, *Proc MM 2004*, 72-79.
- [19] Warne, S., The Mandala Virtual World System, or, Virtual Reality-no strings attached, *Virtual Reality World*, Vol. 2, No. 2, March-April 1994, 65-71.